

University of Groningen

Ultralarge lotteries

Wenmackers, Sylvia

Published in:
Journal of Applied Logic

DOI:
[10.1016/j.jal.2013.03.005](https://doi.org/10.1016/j.jal.2013.03.005)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2013

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Wenmackers, S. (2013). Ultralarge lotteries: Analyzing the Lottery Paradox using non-standard analysis. *Journal of Applied Logic*, 11(4), 452-467. <https://doi.org/10.1016/j.jal.2013.03.005>

Copyright

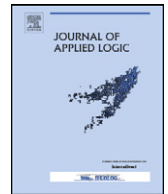
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Ultralarge lotteries: Analyzing the Lottery Paradox using non-standard analysis



Sylvia Wenmackers

Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52, 9712 GL Groningen, The Netherlands

ARTICLE INFO

Article history:
Available online 15 March 2013

Keywords:
Epistemology
Lottery Paradox
Rational belief
Vagueness
Non-standard analysis
Relative analysis

ABSTRACT

A popular way to relate probabilistic information to binary rational beliefs is the Lockean Thesis, which is usually formalized in terms of thresholds. This approach seems far from satisfactory: the value of the thresholds is not well-specified and the Lottery Paradox shows that the model violates the Conjunction Principle. We argue that the Lottery Paradox is a symptom of a more fundamental and general problem, shared by all threshold-models that attempt to put an exact border on something that is intrinsically vague. We propose application of the language of relative analysis—a type of non-standard analysis—to formulate a new model for rational belief, called Stratified Belief. This contextualist model seems well-suited to deal with a concept of beliefs based on probabilities ‘sufficiently close to unity’ and satisfies a moderately weakened form of the Conjunction Principle. We also propose an adaptation of the model that is able to deal with beliefs that are less firm than ‘almost certainty’. The adapted version is also of interest for the epistemicist account of vagueness.

© 2013 Elsevier B.V. All rights reserved.

There are no whole truths; all truths are half-truths. It is trying to treat them as whole truths that plays the devil.
[Alfred North Whitehead [39, p. 14]]

1. Introduction

In order to study the epistemology of yes–no beliefs, in particular the conditions for their rational acceptability, in so far as they are based on probabilistic information, we will focus on a simple example of a lottery. Consider a fair lottery with N tickets, exactly one of which will be randomly selected as the winner. This game of chance can simply be described by a uniform probability function, which assigns a probability of $\frac{1}{N}$ to each ticket. Clearly, the description of a fair lottery does not pose a problem within probability theory, but the interplay of probabilities and rational beliefs triggers epistemological questions. We are interested in what is rational to believe for a participant in such a lottery in the case that N is very large and the winning odds of a single ticket are correspondingly very small.

If you own only one ticket in such a large lottery, it may seem rational for you to believe that your ticket will not win. Buying another ticket does not increase your odds very much, so it still may seem rational for you to believe that none of your tickets will win. Suppose that you keep buying tickets, each with a very small probability of winning, and that you keep believing that none of your tickets will win. At some point, you will own all the lottery tickets and thereby you will

E-mail address: s.wenmackers@rug.nl.
URL: <http://www.sylviawenmackers.be/>.

be certain that one of them will win, which contradicts your belief that none of your tickets will win. This is the Lottery Paradox, originally stated by Kyburg [27].¹

The Lottery Paradox occurs when three *prima facie* plausible principles are combined: the Lockean Thesis, the Conjunction Principle, and the Law of Non-Contradiction. This is the first principle:

Lockean Thesis (LT, Informal Version) It is rational to believe a statement if the probability of that statement is sufficiently close to unity.

The second principle, the Conjunction Principle or CP, states that if it is rational to believe two statements, it is also rational to believe their conjunction. The Law of Non-Contradiction or LNC expresses the idea that it is never rational to believe a contradiction. According to Kyburg himself, it is the employed aggregation rule for beliefs, CP, that causes the paradox [27]. Whereas Kyburg's argument that rational belief is not closed under conjunction was supported by Foley [15] and Klein [25], the idea that CP is the cause of the problem is now a minority position. Some doubt the Law of Non-Contradiction [40], but most contemporary authors are more suspicious of the Lockean Thesis. It has been suggested that LT be modified with a defeater-clause. It seems natural to assume that such a defeater can be made mathematically precise, but Douven and Williamson show that any formally precise defeater does not work to avoid the Lottery Paradox [13], reducing much of the initial appeal of this solution.

In this article, we will analyze the Lottery Paradox as an instantiation of vagueness. After all, the problem only occurs for a lottery ranging over a *large enough* number of tickets, making the probability of winning with a single ticket *small enough*. Also in the informal phrasing of LT, a vague element is present, where it states that the probability has to be *sufficiently close* to unity.

It is the goal of this paper to find a formal solution to the Lottery Paradox that does justice to this vagueness. It may not seem likely that a formal solution of this type exists: what mathematical method can help us out if the problem is intrinsically vague?

We propose to apply relative or stratified analysis [23,24], a type of non-standard analysis. Based on stratified analysis, we will give a formalization of LT and refer to the resulting type of rational belief as 'Stratified Belief'. As it turns out, CP will have to be adapted too, in order to be compatible with this soritic version of LT.

Regarding CP, the conclusion of this paper is close to the position of Kyburg, Foley, and Klein: we find that the Conjunction Principle is too strong to be expected to hold for rational beliefs. However, we do argue in favor of a weakened form of CP. So, like Kyburg, we claim that you would be wrong to keep believing that none of your tickets will win: the repeated addition of an extra ticket with a small probability does not guarantee that the total probability of all the tickets that you own remains small. The total probability of winning will be considerable before you have bought all the tickets. Yet, knowing this does not tell you exactly when you should stop adding tickets or change your opinion. The question "When do the winning odds of a number of tickets cease to be small?" is not all that different from "When does a number of lottery tickets start to be a heap?" In the application of the aggregation rule, we see that induction fails at some point, making the property of rational acceptability of beliefs intransitive.²

In the context of the philosophy of probability, two varieties of probability are considered: objective (or physical) probability on the one hand, and subjective (or epistemic) probability on the other. The probabilities occurring in physics are taken to be objective³ and are thought of as real numbers in the $[0, 1]$ -interval. Subjective probabilities are often referred to as 'degrees of belief' in the Bayesian literature [42,16]. Unlike objective probabilities, degrees of belief do not necessarily have a numerical value. However, in the case of a lottery or other situations in which all relevant information about the objective probabilities is explicitly available, the agent's subjective probability assignments should be equal to the objective probabilities. This requirement has been dubbed "the Principal Principle" by Lewis [31] and we consider it as a minimal, necessary condition for rationality, underlying any attempt to formalize the notion of rational belief. Throughout this article, we will focus on the case in which the subjective probabilities are indeed equal to the objective ones. Alternatively, with Gaifman [17], we may consider a wide spectrum of forms of probability ranging from the purely objective to the purely subjective case, and observe that they all derive from a common kernel. In any case, for our present purposes it is sufficient that an agent assumes the lottery to be fair and reasons under this assumption. Therefore, we may use the term 'probability' without further qualification.

Notation Consider an N -ticket lottery, with N some natural number at least equal to 2. Here, we introduce some notation for probabilities of statements concerning such a lottery. Denote the set of N tickets as: $T_N = \{t_1, \dots, t_N\}$. From this set, exactly one ticket will be randomly selected and assigned to be the winning ticket. If A is a subset of T_N , let $\varphi(A)$ denote the statement that one of the tickets in A is the winner. We introduce a similar notation for loss statements: if B is a subset of T_N , let $\psi(B)$ denote the statement that none of the tickets in B is the winner. Clearly, $\psi(B)$ is equivalent to $\varphi(T_N - B)$.

The assignment of probabilities (P) to win and loss statements can be done as follows:

¹ The paradox can be restated in terms of knowledge [35]. However, here we will address only the original phrasing in terms of rational belief.

² Intransitivity is a typical symptom of problems that are soritic in nature.

³ Of course, even those probabilities are subjective to a certain extent: probability is a way to model a system about which we have insufficient information to predict its behavior with certainty or to summarize information about large numbers of particles.

$$\begin{aligned}
 P(\varphi(A)) &= \#(A)/N \\
 P(\psi(A)) &= 1 - \#(A)/N
 \end{aligned}
 \tag{1}$$

where $\#(A)$ denotes the number of tickets in the set A .

Eq. (1) is not intended to define probability measures (in particular, $P \circ \psi$ is not additive, so it cannot be a measure), but rather to introduce some shorthand notation. If A is a singleton, say $\{t_i\}$, then we write $\varphi(A)$ as φ_i ; we define ψ_i analogously. So, we write $P(\varphi_i)$ as shorthand for $P(\varphi(\{t_i\}))$ and $P(\psi_i)$ for $P(\psi(\{t_i\}))$.

Structure The paper is structured as follows. In Section 2, we frame the Lottery Paradox in the broader context of finding a way to relate real-valued probabilities to binary belief states. We specify three desiderata required for the conversion from probabilities to beliefs. Because the threshold-based model is a popular approach, in Section 3 we review this model and show that it does not satisfy CP. In Section 4, we show that statements about a lottery, for which the paradox may be invoked, show typical traits of vagueness. We claim that the Lottery Paradox occurs in the threshold-based model, precisely because the approach does not deal well with these soritic aspects. In Section 5, we argue in favor of applying ideas from relative analysis to the epistemology of large lotteries. This leads to the main result of this paper: our model of ‘Stratified Belief’ in Section 6. In Section 7, we discuss the relation of our proposed solution to contextualism and the epistemicist account of vagueness. We summarize our findings in Section 8.

2. Mapping $[0, 1]$ onto $\{0, 1\}$

The truth-values of two-valued logic are often represented by 0 and 1. A classical probability function, in the sense of Kolmogorov [26], takes values on the $[0, 1]$ -interval of the real numbers. It has often been suggested that there should be a connection between the two cases. For probabilistic cases with only a finite number of possible outcomes (finite sample space) the equivalence of ‘probability zero’ with ‘logically impossible’ on the one hand, and of ‘probability unity’ with ‘logically necessary’ on the other hand seems warranted. For infinite sample spaces, however, this connection becomes loose: ‘logically impossible’ (or ‘logically necessary’) still implies ‘probability zero’ (or ‘probability unity’, resp.), but the implication of the opposite direction does no longer hold. This connection can be reestablished in the infinite case by using the $[0, 1]$ -interval of a non-Archimedean field as the range of the probability function, instead of the unit interval of the real numbers [5,46].

Also outright belief and yes–no decisions can be considered as two-valued. Thus, underlying the Lottery Paradox, there is a more general question: how do we relate probabilistic information, represented by real numbers in the $[0, 1]$ interval, to simple yes–no judgments (beliefs), which can be represented by the binary values $\{0, 1\}$? Note that there is an asymmetry between i) either you believe p or you don’t, and ii) either you believe p or you believe its negation. Here, the binary values refer to the first interpretation.⁴ So, in accordance with Leitgeb [30], we take beliefs to distinguish between three states: belief, disbelief, and suspension of judgment.

A first answer to the above question could be: we shouldn’t (convert probabilities to unqualified beliefs). If we have detailed information in the form of probabilities, we should stick to that. Indeed, in the lottery case, it is easy to calculate the winning odds of any set of tickets simply by adding the individual probabilities. However, stating that the winning odds of a subset is 0.125, for instance, does not answer the question “Do you believe that all tickets in this set will lose?” There still seems to be a need for a translation.

Foley has argued that it is indeed indispensable to have some way of extracting simple true–false judgments out of detailed probabilistic information [16]. He comes up with some very convincing examples: in court, for instance, the judge or jury has to choose between ‘guilty’ and ‘not guilty’, no matter how fine-grained the information on which they must base their conclusion. In daily life, simple answers are often required to facilitate communication. As Mencken put it: “The public ... demands certainties; it must be told definitely and a bit raucously that this is true and that is false. But there are no certainties” [34, p. 46].

This situation is very analogous to that in image analysis, where black-and-white pictures are sometimes more useful than gray-scale ones. Even the reasons behind this are remarkably similar: conversion of gray-scale images to black-and-white images helps to make certain features stand out more, or to facilitate sending the file by e-mail—think of real-time images sent by a distant space craft—both of which can be seen as facilitating communication. We will encounter this analogy again in the next section.

At this point, we formulate the main requirements for dealing with the relation between probabilities and beliefs:

Desideratum 1 There should be a method to translate continuous probabilities into discrete beliefs.

Desideratum 2 There should be a rule to aggregate these beliefs.

⁴ It is clear that the second notion of belief is even more restrictive: it does not allow a representation of agnosticism, doubt, or suspension of belief, all of which would at least require either a third value besides 0 and 1 or the option not to assign any value at all. We will not deal with this issue here, because adding values shifts but does not remove the problem we want to study: under which conditions is it rational to believe p ? We will take into account different strengths of belief in Section 7.3.

Desideratum 3 The translation method and aggregation rule should be chosen such that together they do not lead to a Lottery Paradox.

First, we look at a candidate for Desideratum 1: a popular model that achieves the translation requirement is the threshold-based model for rational belief; it will be discussed in detail in the next section. Now, let us think of a rule that could be used for Desideratum 2. For logical truths, the aggregation rule is simple: if you start out with two or more true propositions, their conjunction is true as well; this rule of inference is called *Adjunction* or *Conjunction Introduction*. The *Conjunction Principle* (CP) states that something very similar holds for rational beliefs instead of logical truths: if two or more propositions are rationally acceptable, their conjunction is rationally acceptable as well. We can show however (in Section 3.2), that using the threshold-based model for rational belief as Desideratum 1 and CP as Desideratum 2 makes it impossible to meet Desideratum 3. This is no new result: this is precisely the Lottery Paradox.

The approach of this article is to critically examine the threshold-model, and suggest an alternative model without explicit thresholds, which avoids the Lottery Paradox without abandoning CP completely. In Section 6.1, we will propose a new option for Desideratum 1. In Section 6.3, we will find a matching choice for Desideratum 2, such that also Desideratum 3 is fulfilled.

3. Threshold-based model of the Lottery Paradox

In this section we give an overview and critical examination of the wide-spread threshold-based model for rational belief. It rephrases the Lockean Thesis as follows:

Threshold Belief (Informal Version) It is rational to believe a certain statement if the probability of that statement is at least equal to a given threshold value.

This invites a crucial question: what is the value of this threshold used by actual people (in a given context)? And, what ought it be? Usually, the threshold is taken to be a number close to unity, such as 0.999. Achinstein considers a threshold of exactly $\frac{1}{2}$ to be an option, although he prefers the more liberal but less precise condition that the threshold be larger than $\frac{1}{2}$ [1]. He also mentions the less liberal but even more vague condition that the threshold be ‘much larger than’ $\frac{1}{2}$.

When we know that a statement has a sufficiently high probability, Threshold Belief tells us that it is rational to believe the statement fully; this transition is analogous to image processing, where a gray-scale picture can be converted into a black-and-white image by setting a threshold on the brightness: darker pixels become black and brighter pixels become white. In that context too, the question of finding the ‘right’ thresholds is a non-trivial one: in some situations the threshold and even the whole scale may be relative, which leads to adaptive thresholding and image enhancement, respectively [44]. Similar to adaptive thresholding, if the threshold-model is any good at all, it should allow for context-dependent thresholds. Image enhancement is analogous to dramatization: even in one context, it seems as though people use different thresholds in order to contrast cases. This may have little to do with rational beliefs, but if you take into account what it takes for humans to get a message across or remember it [28], context-dependent thresholds definitely serve a function.

Many related questions could be raised, for instance concerning experimental accessibility and measurement precision of the thresholds [12], possible hysteresis [14] etc. However, we will not dwell on these points, for there are more substantial problems with this approach, which will cause us to discard it completely.

3.1. Formalizing the Lockean Thesis

We write $B(x)$ to denote the belief in statement x , and $B(x) \in R_\alpha$ to denote that it is rational for agent α to believe x . Then, the threshold-version of the Lockean Thesis can be formalized as follows.

Threshold Belief (Formal Version)

$$B(x) \in R_\alpha \Leftrightarrow P(x) \geq \theta \quad (2)$$

where θ is the threshold value or cut-off (a real value in $[0, 1]$).⁵

We could have used an agent’s personal probability estimate P_α , but since we assume that all agents use the same probability function for a fair lottery (cf. the aforementioned Principal Principle), we have omitted the subscript.

⁵ Clearly, the range for θ can be made smaller. The interval given here is chosen such that it is broad enough to encompass all the threshold values one may want to consider. There are good reasons to doubt that $\theta \leq \frac{1}{2}$ or $\theta = 1$ are viable choices [1]. Moreover, Achinstein and many other authors suspect that θ should take a value much closer to 1 than to $\frac{1}{2}$.

3.2. Illustration of the failure of CP

We can easily show that:

$$B(\psi_1) \in R_\alpha \text{ and } B(\psi_2) \in R_\alpha \not\Rightarrow B(\psi_1 \wedge \psi_2) \in R_\alpha \quad (3)$$

This can be seen as follows. Applying Eq. (2) to the left-hand side:

$$\begin{aligned} B(\psi_1) \in R_\alpha &\Rightarrow 1 - \frac{1}{N} = P(\psi_1) \geq \theta \\ B(\psi_2) \in R_\alpha &\Rightarrow 1 - \frac{1}{N} = P(\psi_2) \geq \theta \end{aligned}$$

Applying Eq. (2) to the right-hand side:

$$B(\psi_1 \wedge \psi_2) \in R_\alpha \Rightarrow 1 - \frac{2}{N} = P(\psi_1 \wedge \psi_2) \geq \theta$$

However, $1 - \frac{1}{N} > 1 - \frac{2}{N}$; thus $1 - \frac{1}{N} \geq \theta$ does not guarantee that $1 - \frac{2}{N} \geq \theta$ (unless when $\theta = 1$, but this case is of little interest for modeling belief anyway).

In words, this shows that, given a value for the threshold, θ , there exists a sufficiently large lottery such that the losing odds of any specific ticket are at least equal to the threshold, but the losing odds of two tickets are below it. If rational acceptability is identified with having probability above a certain threshold—that is to say, the ‘if’ in the informal phrasing of LT is interpreted as ‘only if’—, the conjunction of only two rational beliefs may fail to be rationally acceptable.

4. The vague lottery: a heap of tickets

In this section, we step away from the existing threshold-based model of LT and take a fresh look at the Lottery Paradox. The first three observations of this section will fuel our search for a new formalization of LT, which will be the foundation of a new solution to the Lottery Paradox. The fourth observation foreshadows the idea that also CP will have to be adapted.

4.1. Observation 1: The soritic lottery

In formal epistemology, the Lottery Paradox may be seen as an easier (clearer) problem to work on as opposed to cases in which the probabilities are not explicit, such as the ‘Preface Paradox’ [33] that deals with a book containing a lot of statements, in each of which the author has ‘very high’ confidence. This viewpoint may be misleading: to invoke the Lottery Paradox, the number of tickets N just has to be ‘large enough’. Another way of saying this is: the probability of winning $\frac{1}{N}$ has to be ‘small enough’. ‘Large enough’ and ‘small enough’ are vague concepts: vagueness is at the heart of the problem (both for the Lottery and the Preface Paradox).⁶

There is also vagueness in the informal version of the Lockean Thesis (LT), which mentions a probability *sufficiently close* to unity, but this vagueness is not reflected in the usual formalization of LT with thresholds. The threshold-model has stepped into the trap of illusory exactness: it attempts to set a sharp boundary around what is large, in particular what is a large enough probability. The idea that ‘being large’ is a well-defined property does not agree with our normal use of the concept. It is the position of this article that counter-intuitive results such as the Lottery Paradox follow from just this property. The Lottery Paradox is merely a symptom of this deeper problem.

This first observation suggests that we should formalize LT such that the vague aspect of it is respected and apply this vague version of LT to the lottery case. Now, the puzzle is how to deal formally with the kind of vagueness at issue here.

4.2. Observation 2: Contextual element

Yet another formulation of the lottery case, which although still vague, at least has the advantage of making clear that ‘large’ and ‘small’ are relative concepts: in order to make it plausible that it is rational to believe that none of your tickets will win, the number of tickets that you own has to be very small compared to the total number of tickets. Although it is less quantitative, in a way the latter statement is more informative than “You own two tickets in a fair 10 000 000-ticket lottery with one winner”.

The observation that the vagueness involved in the Lottery Paradox is relative suggests that we should allow a contextualist element in the solution.

⁶ Gaifman claims that there is an “essential difference” between the Sorites and the Lottery Paradox, in particular, that there is “nothing semantic” about the Lottery Paradox [19, p. 15]. On my view, however, there is an essential soritic aspect to the Lottery Paradox.

4.3. Observation 3: Analogy with fair lottery on \mathbb{N}

As Lavine points out [29, p. 389]: “It is a familiar idea that our knowledge about the infinite is obtained by in some sense extrapolating or idealizing our knowledge of and about the finite.” This idea can also be applied in the opposite direction: large finite phenomena are often modeled by infinite ones. In physics, a long thin cylinder may be taken to be an infinitely long wire, a large surface to be an infinitely large area (e.g. by using periodic boundary conditions), and so on. For our problem, it seems natural that a *sufficiently large* lottery behaves qualitatively the same way as does an *infinite* lottery.

If we want to describe a lottery on \mathbb{N} , it turns out that we do not have the freedom to assign fair odds to the tickets within the classical axiomatization of probability theory developed by Kolmogorov [26]. A fair lottery on \mathbb{N} can be described if we drop the requirement of Normalization or that of Countable Additivity. If we, like de Finetti [8], choose the latter option, the probability of winning of a single ticket (or any other finite subset of \mathbb{N}) is zero, the same probability as we assign to the impossible event. Real-valued, finitely additive probabilities are not fine-grained enough to distinguish between the impossible event and some possible but ‘highly unlikely’ events. Within the real numbers, there is no way of quantifying just how unlikely these events are; all we can express is that their probability is zero. Although there is a clear *qualitative* difference between possible and impossible events, some of the associated probabilities are *quantified* by the same number.⁷ Clearly, this situation is unsatisfactory from an epistemological point of view. If we allow infinitesimals in the range of the probability function, it is possible to distinguish between the probability of the empty set and a non-empty set (cf. [5,46]).

Thus, the infinite version of our problem requires the use of non-standard analysis (NSA), originally developed by Robinson [43].⁸ The idea of NSA is to extend \mathbb{N} and \mathbb{R} to the strictly larger sets ${}^*\mathbb{N}$ and ${}^*\mathbb{R}$. ${}^*\mathbb{N}$ is called the set of hypernaturals and contains infinite numbers, strictly larger than any natural number. ${}^*\mathbb{R}$ is called the set of hyperreals, and next to infinite numbers it also contains their inverse—infinitesimals—which are smaller in absolute value than any strictly positive real.

If we consider a fair lottery on a finite subset of \mathbb{N} , the probability assignment is unproblematic: the winning odds of any non-empty subset are non-zero. However, in our reasoning about such a lottery we sometimes deal with very small probabilities as if they were zero. In such a situation, our beliefs are not fine-grained enough to distinguish between the impossible event and some possible but ‘highly unlikely’ events. Because of the analogy between the infinite lottery puzzle and the Lottery Paradox, we ask: can NSA be applied to solve the latter problem too?

It should come as no surprise that the study of an infinite lottery leads to a system in which one has infinite numbers and infinitesimals available. However, at first sight, there seems no good reason to bother with infinities when dealing with beliefs about a finite lottery, no matter how large. Although we do not want to introduce infinite numbers, this is not a valid objection to the use of NSA for this problem, since there are approaches to NSA that do not extend the standard sets into the transfinite, but work entirely within the standard sets \mathbb{N} and \mathbb{R} . This is true for Nelson’s ‘internal set theory’ (IST) [36], as well as Hrbacek’s relative analysis (inspired on IST) [23].

A promising aspect of relative analysis is that it is able to cope with soritic concepts. Some vagueness is present in Robinson-style NSA, too. To see this, observe that the standard sets \mathbb{N} and \mathbb{R} have no largest element. Likewise, in ${}^*\mathbb{N}$ and ${}^*\mathbb{R}$ there is no smallest infinite number: although for any given number, it is easy to determine whether it is finite or infinite, the border between the finite numbers and infinite numbers cannot be pinpointed by giving a number at or near it. In ${}^*\mathbb{R}$, there is no largest infinitesimal, nor a smallest positive non-infinitesimal number. Stated informally, the result of the internal-set approach is that the vague border between the finite and infinite gets ‘pushed down’ into a vague border between standard and large size, whereas that between finite and infinitesimal ‘scales up’ into a vague distinction between standard and small size. With relative analysis, we will model the small probabilities that are indistinguishable from zero as ‘relative infinitesimals’ or ‘ultrasmall numbers’.

4.4. Observation 4: Weakening CP

For statements supported by probabilistic considerations, it is not self-evident that CP should hold. First, let us recall how to calculate the probability of a conjunction. In the case of two tickets for the same lottery, their losses are dependent events and the probability of a loss statement decreases linearly with the number of tickets.⁹ Nevertheless, if we start out believing a statement that has a sufficiently high probability, the aggregation of some additional statements with equally high probability will not dramatically change our (degree of) belief in the conjunction. So we do not expect to see CP fail completely either. For the aggregation of beliefs, we expect to find a weakened version of CP. In particular, for beliefs concerning a large lottery, what we expect intuitively is this¹⁰:

⁷ The problem is not limited to impossible and highly unlikely events: there are many more events that are qualitatively distinguishable (they differ up to a finite subset of \mathbb{N}) but get exactly the same probability assignment. In particular, probability one is not only assigned to the necessary event (‘Some ticket of \mathbb{N} will win’), but also to infinitely many other events: \mathbb{N} minus any finite subset.

⁸ Although the phrase ‘non-standard’ in the name may have a negative connotation (suggesting that the theory is abnormal), it comes from the notion of a ‘non-standard model’, which is an unintended interpretation of a formal system. These non-standard models have turned out to be very useful in many areas of mathematics, including that of analysis. See also [18].

⁹ Consider two different tickets i and j in an N -ticket lottery. Then, the event that ticket i will not win has probability $P(\psi_i) = 1 - \frac{1}{N}$, while the event that both tickets i and j will not win has probability $P(\psi_i \wedge \psi_j) = 1 - \frac{2}{N}$.

¹⁰ A similar observation is made by Gaifman [19, p. 16]: “[A] derivation should not use too many Sorites conditionals, where ‘too many’ means that the set of conditionals makes it possible to traverse a Sorites chain”.

Weakened Conjunction Principle for Beliefs (Informal Version) It is acceptable to aggregate a few rationally acceptable beliefs, but the conjunction of many rationally acceptable beliefs is not necessarily rationally acceptable.

In a threshold-based model, this intuitive rule is violated, for we cannot even allow the conjunction of two beliefs, as we have seen in Section 3.2.

Another way of seeing that we may have to weaken CP is by observing the following statements:

- The probability of winning for one ticket is small.
- The probability of winning for two tickets is small.
- ...

In mathematics, it is often tricky to correctly interpret a continuation with an ellipsis: it suggests a type of limit process. (The limit of classical calculus is not the only option; we will come back to this.) Here, the same care is needed, for the ellipsis does not generalize to the (incorrect) statement: “The probability of winning for all tickets together is small”, but rather to:

- The probability of winning for a few tickets is small.

Whereas ‘all tickets’ refers to a definite number, ‘few’ is a vague term of course. As we have already seen in Observation 3, relative analysis may be able to deal with this formally.

5. Introduction to relative analysis

By combining Observations 1, 2, and 3 from previous section, we know that we need a formalism that is capable of dealing with vagueness, provides contextual elements, and is a form of NSA. The good news is that we do not have to develop a theory from scratch: there is a formal system available that precisely formalizes the contextual and vague concepts of largeness and smallness, and more generally different scales of magnitude. It is an approach to NSA developed by Hrbacek [23,24], called ‘relative analysis’ or ‘stratified analysis’.¹¹ In this section, we introduce the approach of relative analysis, in order to apply it to the Lottery Paradox in the next section.

5.1. Scales of magnitude

Because stratified analysis was developed to deal with soritic quantities, let us apply it to an example dealing with a heap, or rather a bucket, of sand. Relative to the bucket, which is mesoscopic (neither small nor large), a single grain of sand is negligibly small (microscopic). The whole beach, however, is gigantic (macroscopic) compared to the bucket. This is depicted schematically in Fig. 1.

The numbers $0, 1, 2, 3, \dots$ and $\frac{1}{2}, \frac{1}{3}, \dots$ are mesoscopic numbers in relative analysis. They are observable on the coarsest context level (to be defined below). In order to apply relative analysis, we may let the numbers correspond to a physical quantity. A bucket of sand is a conceivable quantity to us, humans: it is one “dose” of sand, so we may use it as a unit of sand (much like dm^3 is a unit of volume). In a different application (e.g. when describing the viewpoint of a small animal on the beach), it may be more useful to take one grain of sand as the unit (or mm^3 as the unit of volume). In this application, the grain of sand is mesoscopic, while both the bucket and the beach are macroscopic. Thus, one and the same quantity can be both gigantic and negligibly small, depending on the choice of unit (i.e. which physical interpretation is given to the number 1, which is always considered to be mesoscopic in relative analysis).

This example indicates that our observations are related to different scales. One aspect of this can be understood from Lavine’s approach to finite set theory [29]. He uses a similar example (albeit with beans instead of grains of sand) as a physical model for learning addition. Although the bucket contains a finite number of grains, we have no idea how many: the bucket is an ‘indefinitely large’ supply of grains of sand. Likewise, the beach is indefinitely large compared to the bucket. Lavine points out that whenever necessary, we may take a larger supply of sand, but this does not mean we ever need an infinite amount. The context of set theory is less convenient to discuss ‘indefinitely small’: whereas we can consider larger and larger collections, at the bottom we find singletons, and cannot look at smaller scales. In other words, we cannot represent scales (of largeness and smallness) well by the counting numbers (\mathbb{N}) alone, but since the real numbers are closed under inversion, \mathbb{R} may do better.¹²

¹¹ Of course, there are other formal systems that deal well with vagueness and imprecise probabilities, such as fuzzy logic, which has been applied to belief by Booth and Richter [6], and interval-valued probabilities of Dempster–Shafer theory [9], but we will not discuss those further, here.

¹² Actually, \mathbb{Q} would suffice for our current purpose.

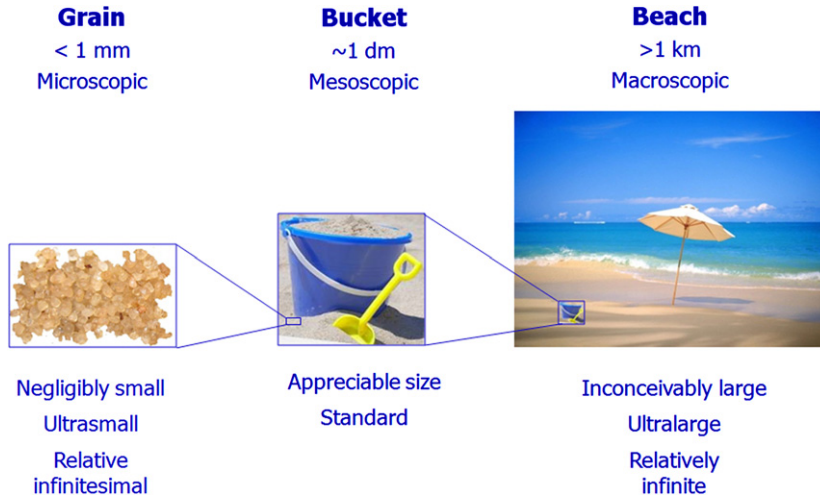


Fig. 1. A typical grain of sand is less than 1 mm in size. This is negligibly small compared to a typical bucket. The length of the beach is much larger than the dimensions of a bucket, so much that the factor is inconceivable to us.

5.2. Levels

Relative analysis formalizes the aforementioned scales of magnitude as levels, a new concept intended to correspond with the intrinsically vague concept of largeness. As was mentioned already, unlike Robinson-style NSA, which extends \mathbb{N} and \mathbb{R} to the strictly larger sets ${}^*\mathbb{N}$ and ${}^*\mathbb{R}$, relative analysis works within \mathbb{N} and \mathbb{R} . To obtain the concepts of infinitesimals and infinite numbers within \mathbb{R} , relative analysis adds the new primitive binary predicate \sqsubseteq to the language [23].¹³ The meaning of this predicate is fixed by axioms, but informally $x \sqsubseteq y$ means that the number x is observable at every level where the number y is observable: the number on the left is observable at the coarsest level where that at the right is observable, and possibly at coarser levels too. To express that x and y are observable at the same level, we may write $x \sqsubseteq y$ and $y \sqsubseteq x$. To simplify the formalism, the axioms can be represented in terms of the aforementioned levels. So, “ x belongs to the level of y ” replaces $x \sqsubseteq y$. Levels stratify the set \mathbb{R} into different levels or scales of magnitude; therefore, relative analysis is also referred to as stratified analysis.

5.2.1. Axioms for levels

In [24], there are eight axioms that fix the meaning of the level-concept. We paraphrase them here:

1. For every finite collection of real numbers, there is a coarsest level at which all the specified numbers are observable.
2. There are also finer levels, on which more numbers are available. Two levels can always be compared: we can always say which level is *at least as fine as* the other. Although levels are not sets (see Section 5.4), $V_1 \supseteq V_2$ is used to indicate that level V_1 is at least as fine as level V_2 .
3. For any level, there exist non-zero real numbers that are ultrasmall compared to it.
4. Neighbor Principle: at every level there is a best approximation to any real number that is not ultralarge relative to that level.
5. Closure Principle: any number, function, operation, or set that is defined without mention of levels from parameters that are observable at a certain level is itself observable at that very same level. This level is called the context level or observation level.
6. Stability Principle: if a statement is true about its context level, then it also holds for any finer level. Using Stability, Closure can be generalized to the Transfer Principle, an indispensable tool in any approach to NSA [4].
7. Definition Principle: for any internal statement (i.e. any statement that makes no reference to levels coarser than the context level) and any set A of real numbers, there exists a set B whose elements are exactly those elements of A for which the internal statement holds.
8. Density of levels: given two levels $V_1 \subset V_3$, there is a level V_2 such that $V_1 \subset V_2 \subset V_3$.

¹³ In Nelson's Internal Set Theory, a new *unary* predicate is introduced to signify “ x is standard”. To allow for multiple levels of standardness, Péraire and Wallet introduced a *binary* predicate to signify “ x is standard compared to y ”; they called their theory ‘relative internal set theory’ (RIST) [38]. Also Hrbacek's relative analysis is based on this binary predicate, for which he uses the symbol \sqsubseteq .

In relative analysis, the number 1 is always standard at every level. From the viewpoint of applications, this poses no limitation: after choosing any non-zero number as the unit of interest, one can always divide the whole scale by this value and thus achieve normalization. This points out the fractal-like structure of the real line: it is self-similar on all scales.¹⁴

5.2.2. Further terminology

Let us introduce some more of the vocabulary of relative analysis. Some numbers can be used to define a level V (cf. axiom 1): these numbers are called ‘standard’ compared to level V ; they are also said to be ‘observable at that level’.

Given a level V , there are non-zero numbers ‘ultrasmall compared to’ other numbers (\ll_V); ultrasmall numbers are ‘relative infinitesimals’. Likewise, on level V there are numbers ‘ultralarge compared to’ other numbers (\gg_V , ‘relatively infinite numbers’). A number that is not ultralarge compared to level V is called ‘finite’ compared to this level. Zero is the only number that can belong to a level while being ultrasmall compared to that level. A level does not contain any numbers that are infinite in comparison to it.

If the difference between two numbers is ultrasmall compared to level V , the numbers are ‘ultraclose’ to each other on level V (\simeq_V). In other words, they are indistinguishable (on that level), for their difference is negligible (on that level).

5.3. What levels are: a predicate on the domain

Like Lavine’s indefinitely large sets within \mathbb{N} , levels on \mathbb{R} are predicates on the domain. Thus, a level is a collection of real numbers. It contains all the numbers that have a unique name. At any given point in time, there can only be finitely many numbers that have a unique name. Therefore, the set of reals always contains infinitely many numbers that are larger than any uniquely named number (cf. Lavine’s indefinitely large numbers), as well as infinitely many that are smaller than any uniquely named real number. Thus there is always ‘room’ for ultralarge and ultrasmall numbers. (Of course you can refer to these ultralarge or ultrasmall numbers indirectly, but you can never give all of them a unique name in finite time.)

When we *apply* relative analysis (as opposed to making a contribution to the development of its mathematical formalism), we may allow different criteria for numbers to be standard—not just uniquely named ones. Some examples:

- Distances in an image that are above the lower detection limit (resolution) and smaller than the upper detection limit (field of view).
- The quantity of sand measured as numbers of buckets, including partially filled buckets (fractions).
- Numbers of lottery tickets and their probabilities of winning.

5.4. What levels are not: sets

It is crucial to note that levels are not sets; in particular, the principle of mathematical induction does not hold for levels. For instance, adding or multiplying two numbers that are observable on a level V will result in a number that is standard compared to the level as well. The addition and multiplication can be generalized, but the result is only guaranteed to be standard for a relatively small (standard) number of terms or factors.

The set of real numbers, \mathbb{R} , forms a complete metric space, which means that every Cauchy sequence of real numbers converges to a real number. Unlike \mathbb{R} , the extended set $^*\mathbb{R}$ is incomplete and as a consequence the limit concept of classical calculus is not available in it. However, within the various approaches to NSA, there is a different limit operation available (such as the alpha-limit in Alpha Theory [3]).

Internal set theory and the related approach of Hrbacek do not introduce the non-standard extension $^*\mathbb{R}$, but work on the standard set \mathbb{R} . Yet, there is a form of incompleteness that occurs within these approaches too: they introduce new predicates that do not necessarily have a defined limit. Mathematical induction is like a limit operation on sets; as noted at the beginning of this paragraph, it does not apply to levels. This elucidates the meaning of the ellipsis following the sentences “The probability of winning of one ticket is small”, “The probability of winning of two tickets is small”, ... Just as there is no last finite number in $^*\mathbb{N}$, there is no sharp boundary around the number of tickets which collectively have a small probability of winning. Whereas the (incorrect) application of induction would lead us to a crisp, but false conclusion (“The probability that the set of *all* tickets contains the winner is small”), the correct interpretation using levels leads to a sound generalization, albeit one containing a vague term (“The probability that the set of *a few* tickets contains the winner is small”).

We may rephrase the analogy between extension-style NSA and relative analysis as follows. Asking what is the supremum of a set in $^*\mathbb{R}$ does not always make sense. Because $^*\mathbb{R}$ is incomplete, there is not necessarily an element of $^*\mathbb{R}$ at the point to which the limit seems to converge. The question posed may well point to a gap. Asking what is the supremum of a level in \mathbb{R} is equally misguided. There is no such element; the question points to a gap in the system. There is no largest real number that is standard on a given level. The borders of a level are vague.

¹⁴ With relative analysis, the fractal-like structure of \mathbb{R} comes to expression in the fact that, for any level V , true formulas that quantify only over levels finer than V are exactly the same. However, not all formulas hold for all levels: for instance, a formula dealing with the existence of coarser levels will not hold at the coarsest level.

Some form of incompleteness—the presence of gaps in the set itself or gaps related to new predicates on the set—is necessary to invoke infinitesimals.¹⁵ If you try to make the vague borders of a level precise, relative analysis collapses to standard analysis. This points out that a model of beliefs based on relative analysis is incompatible with a model involving explicit, sharp thresholds.

5.5. The grain of sand, the bucket, and the beach

The sand example given at the beginning of Section 5.1 can now be represented as: Grain of sand \ll_V Bucket (standard) \ll_V Beach, where the level V is that of a child playing on the beach (where V also contains the bucket).¹⁶

It is very likely that it does not make any difference to the child exactly how many grains of sand a bucket contains. If there were one grain more or less, he would not notice it, so he is indifferent to that. The same holds for any small number of grains that are added or removed, where ‘small’ means ultrasmall compared to the total number of grains of sand in the bucket. Clearly, the child will notice (and may care) if, for instance, all or half of the grains of sand in his bucket are removed.

The transition between will and will not notice is vague. In an experimental setting, this could mean that on one occasion the child will notice a certain difference, whereas he might not register this at another time. If you keep removing grains of sand one by one, you may get further than if you remove them all at once. We will not dwell on these general properties of vagueness, for there is another aspect that is more relevant to the application that we have in mind.

The child may weigh the bucket to estimate the number of grains of sand in it, or may even count the actual number of grains of sand.¹⁷ In the first case, he will be able to tell the difference between at least some buckets that appear equally filled to the unaided eye. His increased ability to distinguish quantities can be represented by a finer level. In the second case, given enough time the child will be able to spot *any* difference in the filling of the buckets: it is at the finest level relevant for this example.¹⁸

6. Analysis of the Lottery Paradox using relative analysis

We are now ready to formulate a new definition for rational belief. This new definition is inspired by the Lockean Thesis, but developed along the lines of stratified analysis rather than a view based on thresholds.

6.1. Stratified model for rational belief

Whereas the sand example dealt with our sensory limitations (the (in-)ability to perceive differences in heaps of sand), forming rational beliefs about a lottery has to do with our mental ability to deal with numbers. Here, the problem is not that we do not see the difference between 0.999 and 1, but rather that we do not always attach any significance to the difference. Sometimes we deal with 0.999 as we would with 1, while in other contexts we may consider the difference as highly relevant. Even when we know probabilities are quantitatively different, our means of categorizing them qualitatively (large/small, comparable/different, ...) are limited. This limitation is useful, for we are finite beings with finite capacities.

Epistemology often deals with highly idealized agents, but in order to make sense of the Lottery Paradox, it is important to take into account at least this descriptive element: humans use scales of magnitude to make qualitative assessments. Given this limitation, we ask how we can deal with this as well as possible in the formation of rational beliefs. The use of scales may introduce a type of rounding error, which may go unnoticed when dealing with direct sensory information, but may produce some strange consequences if the starting information is precise and explicitly quantitative, as is the case with judgments based on probabilities.

If we have to judge whether or not we believe a certain event will take place, based on the numerically specified probability of the event, we have to compare the given probability value to unity. If the given value is zero or one, we can immediately answer that we do not or do believe that the event will happen. For intermediate values, we have to judge on a case-by-case basis whether or not the provided probability is close to unity: we will model this as being ‘ultraclose to unity’ in the sense of relative analysis.

¹⁵ Levels are ‘gappy’, just like collections of finitistically acceptable numbers are; see for instance [32, p. 475].

¹⁶ Observe that although a grain of sand is at a *smaller* scale of magnitude than a bucket and a beach is at a *larger* scale of magnitude, both appear at a *finer* level when the bucket is taken to be at the standard level. Thus, although the intuitive notion of scales of magnitude was an important motivation for developing a theory in terms of levels, the terms are not synonymous.

¹⁷ It seems unlikely that the child would do any of this spontaneously, but he might be motivated to do so for a project at school or because there is a prize connected to getting the answer right. Of course, you are free to see this as an allegory of what is going on in science.

¹⁸ This idea of different levels in relative analysis is related to that of ‘degrees of availability’ in finite mathematics [29]. Of course, if the next task were to keep track of the number of atomic nuclei (or even smaller structures) in the bucket, and the child had an electron microscope (and a really long holiday ...) then an even finer level could be appropriate. This is possible, for we are working in the real numbers (a complete set with perfect self-similarity), where we have an unlimited supply of finer (axiom 3) levels available. We may also model this differently: by choosing the smaller structure of interest as the new unit of the context level.

Definition 1 (*Stratified Belief, SB*). It is rational for agent α to believe x on a level V if and only if the probability of x is indistinguishable from unity on the context level of the agent. More formally:

$$B(x) \in R_{\alpha,V} \Leftrightarrow P(x) \simeq_V 1 \quad (4)$$

Compared to the threshold-version of the Lockean Thesis, there is a new contextual element present here—that of a level—which will be discussed further in Section 7.¹⁹ Expressed in words, SB says: relative to a certain level V , it is rational for an agent to believe a proposition if and only if the probability of the proposition is ultraclose to unity as compared to the level. Other ways of formulating this condition for rational belief is that the probability should be indistinguishable from unity (on a given level) or equal to unity up to an infinitesimal (relative to that level).

Also in the context of probabilistic approaches to conditionals, infinitesimals have been applied to interpret conditionals by Adams [2]: the statement “If A then B ” is read as “The conditional probability of B given A is larger than 1 minus an infinitesimal” or symbolically: $P(B|A) > 1 - \epsilon$.²⁰ The latter expression can also be used in default reasoning to represent the statement “Normally, if A then B (but there may be exceptions)”. In the context of Adams’ work, ϵ does not refer to an infinitesimal in the sense of NSA or relative analysis, but to the ϵ, δ -framework of standard analysis. However, the interpretation is quite similar to that of SB: a probability that is infinitesimally close to unity indicates ‘almost certainty’ (which includes full certainty).

Whereas the version of SB given in this section models rational belief as almost certainty, we will relax this condition in Section 7.3.

6.2. Stratified Belief applied to a large lottery

Now we can apply the definition of SB to the case of an (ultra)large lottery. For an N -ticket lottery, where N is large (as judged by a specific agent), consider level V_{lott} which contains 1 but not N .²¹ In other words, on this level 1 is standard and N is ultralarge. In that case, $\frac{1}{N}$ is ultrasmall compared to 1 on this level ($N \gg_{V_{\text{lott}}} 1 \gg_{V_{\text{lott}}} \frac{1}{N}$), which shows that level V_{lott} is a good starting point for discussing the probabilities of an (ultra)large lottery. The standard numbers in the $[0, 1]$ -interval are of particular interest, since, in the current view, they guide the agent in his belief-forming practices.

Since $\frac{1}{N}$ is ultrasmall, $\frac{1}{N} \simeq_{V_{\text{lott}}} 0$, we find that:

$$1 - \frac{1}{N} \simeq_{V_{\text{lott}}} 1$$

Let us consider two groups of tickets: 1) the set of all but one lottery tickets ($N - 1$) and 2) the set of all tickets (N). The above equation formalizes that, on a certain level, an agent is not able to appreciate the difference between the winning odds of both sets of tickets.

We may interpret the same equation in terms of losing odds instead of winning odds: then it expresses that on this level, the agent also cannot distinguish between the losing odds in the situation 1) in which he owns one ticket and 2) in which he has no ticket at all. If the difference between having a ticket or not appears negligible to an agent, it is only rational for him to believe that his single ticket will not win.²²

6.3. Aggregation of stratified beliefs

The next question is whether the Conjunction Principle holds for this type of beliefs. We will see that it is allowed, but only within a level. This Stratified Conjunction Principle (SCP) formalizes our intuitive weakening of CP given in Section 2.

Although we started our search for a solution to the Lottery Paradox from the viewpoint that we need to adapt the formalization of the Lockean Thesis, it does turn out that we have to adapt CP too.

6.3.1. Aggregating beliefs concerning single lottery tickets

First we show that the conjunction of two rational beliefs (in the sense of SB), each concerning only one lottery ticket, amounts to a new rational belief.

¹⁹ Also in the discussion of Threshold Beliefs, it has been remarked that the threshold should be thought of as context-dependent (see e.g. [22, p. 246]); in SB, this consideration is taken into account as an explicit parameter.

²⁰ Thanks to Sonja Smets for the pointer.

²¹ Before proceeding, we should check that such a level exists. Axiom 3 of [24] ensures that there always exists a number ϵ that is ultrasmall compared to the given level; thus, all numbers $N > \frac{1}{\epsilon}$ are sufficiently large for such a level to exist. We can give examples of values for N that are not large enough (such as $N = 2$), but because we cannot give an explicit example of ϵ , there also is no clear threshold for N being large enough—a remark that should sound familiar at this point.

²² Because the winning odds of all but one ticket are equal to the losing odds of one ticket, on the same level they are either both distinguishable from unity or both indistinguishable from it. This may be a drawback if we want to model an optimistic person, who is less worried that he will not win if he owns all but one ticket than he is expecting to win if he owns just one ticket, and vice versa for a more pessimistic mind. Although this could be modeled using different levels for the two cases, or levels that are asymmetric under inversion, it is probably better to regard the psychological factors of risk seeking and risk aversion as outside the scope of what the model of Stratified Belief intends to capture.

Definition 2 (Stratified Conjunction Principle, SCP).

$$B(\psi_1) \in R_{\alpha, V} \text{ and } B(\psi_2) \in R_{\alpha, V} \Rightarrow B(\psi_1 \wedge \psi_2) \in R_{\alpha, V} \quad (5)$$

To show that SCP indeed holds, apply Eq. (4) to the left-hand side:

$$B(\psi_1) \in R_{\alpha, V} \Rightarrow 1 - \frac{1}{N} = P(\psi_1) \simeq_V 1$$

$$B(\psi_2) \in R_{\alpha, V} \Rightarrow 1 - \frac{1}{N} = P(\psi_2) \simeq_V 1$$

and apply Eq. (4) to the right-hand side of SCP:

$$B(\psi_1 \wedge \psi_2) \in R_{\alpha, V} \Rightarrow 1 - \frac{2}{N} = P(\psi_1 \wedge \psi_2) \simeq_V 1$$

Since $1 - \frac{1}{N} \simeq_V 1 - \frac{2}{N}$, $1 - \frac{1}{N} \simeq_V 1$ guarantees that $1 - \frac{2}{N} \simeq_V 1$.

At first sight, SCP takes the same form as CP. To see that SCP is nevertheless a weakened form of CP, note that SCP does *not* generalize to the conjunction of any number of stratified beliefs. In particular, it does not hold that the conjunction of an ultralarge number of rational beliefs is rational. That is to say, $B(\psi_1) \in R_{\alpha, V}$ and ... and $B(\psi_N) \in R_{\alpha, V} \not\Rightarrow B(\psi_1 \wedge \dots \wedge \psi_N) \in R_{\alpha, V}$ (because $1 - \frac{N}{N} = 0 \not\simeq_V 1$).

6.3.2. Generalization of SCP

Although SCP does not generalize to the conjunction of ‘many’ (i.e. an ultralarge number of) and in particular all beliefs, it does allow the conjunction of ‘a few’ (i.e. a standard number of) beliefs, or—which is equivalent—the conjunction of two beliefs, each concerning ‘a few’ (standard number of) tickets. As such, SCP can be considered as the formal counterpart of our intuitive weakening of CP stated in Section 2.23

Here, we will prove this slightly stronger version of SCP, that is valid for arbitrary events E_1 and E_2 , not just singletons²⁴:

$$B(\psi(E_1)) \in R_{\alpha, V} \text{ and } B(\psi(E_2)) \in R_{\alpha, V} \Rightarrow B(\psi(E_1) \wedge \psi(E_2)) \in R_{\alpha, V}$$

The proof of this conjunction rule is as follows: by the definition of SB in Eq. (4), the two assumptions are equivalent to $P(\psi(E_1)) \simeq_V 1$ and $P(\psi(E_2)) \simeq_V 1$, respectively. Thus, $1 - \frac{\#(E_1)}{N} \simeq_V 1$ and $1 - \frac{\#(E_2)}{N} \simeq_V 1$, which imply that $\frac{\#(E_1)}{N} \simeq_V 0$ and $\frac{\#(E_2)}{N} \simeq_V 0$. If two numbers are ultraclose to zero, this is also true for their sum: $\frac{\#(E_1)}{N} + \frac{\#(E_2)}{N} \simeq_V 0$. Since $\#(E_1) + \#(E_2) \geq \#(E_1 \cup E_2)$, we find that $\frac{\#(E_1 \cup E_2)}{N} \simeq_V 0$. This implies that $1 - \frac{\#(E_1 \cup E_2)}{N} \simeq_V 1$, or in terms of probability: $P(\psi(E_1 \cup E_2)) \simeq_V 1$. Because of the definition of ψ , this is equivalent to: $P(\psi(E_1) \wedge \psi(E_2)) \simeq_V 1$. Finally, by the definition of SB, we conclude that: $B(\psi(E_1) \wedge \psi(E_2)) \in R_{\alpha, V}$.

7. Relation to philosophical theories and application to other problems

7.1. Relation to Gaifman’s approach to vagueness

Gaifman proposes a framework for contextual logic [19]. It is based on classical two-valued logic, as it is intended to describe an agent who has to make a yes–no decision; in that sense, it is similar to our model of Stratified Belief. First, Gaifman introduces a modal operator which can be read as ‘definitely’ and represents semantic necessity (without metaphysical or ontological repercussions), the negation of which represents semantic possibility. A third operator, the borderline operator, is defined as neither definitely (the proposition) nor definitely not (the proposition). A nice consequence of this approach is that higher-order vagueness can be modeled as an iteration of the borderline operator. Gaifman observes that “vague predicates are associated with some notion of degree”, and shows that these degrees follow from the above modalities [19].

The main goal of the current paper was to analyze the Lottery Paradox, rather than to propose a formalization of vagueness, and the chosen framework was that of non-standard analysis, rather than that of modal logic. Nevertheless, there seem to be many aspects that are common to the current work and that of Gaifman [19]. It would be nice to explore these similarities in more depth, as they may suggest deeper connections between Gaifman’s contextual logic and Hrbacek’s relative analysis, but such an endeavor goes beyond the scope of the current paper.

²³ The result obtained here is similar to Magidor’s observation that strict finitism (in particular when combined with a cut-free proof system) may allow for Sorites arguments with true premises and a false conclusion [32].

²⁴ Thanks to Karel Hrbacek for suggesting this.

7.2. Relation to contextualism

Skepticism tells us that there is always the possibility that our whole life is an illusion and we are just brains in a vat [41], or, that when we say “We will be there in one hour”, the Earth will get hit by a huge meteor and we never get there. (Similar examples are considered by Harman [21].) In daily conversation, it would be tiresome to always sum up these and similar highly unlikely events. Yet, it seems like we should, for there is no way to exclude these options with absolute certainty. The analysis in terms of levels gives a post-hoc justification of what we actually do: we usually treat the unlikely events as infinitesimals, but there is a finer level available—which may be relevant in scientific or philosophical contexts—on which even these minuscule probabilities are appreciable.²⁵ Moreover, we may compare the highly unlikely events: I judge the probability of my whole life being an illusion as ultrasmall, even compared to the probability of the Earth being hit by a large meteor in the next hour, which is itself ultrasmall (compared to the chance of a die landing with the 6 face up, for instance). Because we are free to use a different level in different situations, we may say to a friend that we believe that we will arrive in 60 minutes because our GPS says so, but deny that we believe that the extinction of the dinosaurs has been caused by anything other than a meteor. In both cases we use the verb “to believe”, which suggests a fixed scale of belief, but apparently the scale does depend on the context.

Contextualism is “a theory according to which the truth-conditions of knowledge-ascribing and knowledge-denying sentences ... vary in certain ways according to the context in which they are uttered” [10, p. 2]. The variable parameter is the epistemic standard. We hope that it is clear at this point that a rise in epistemic standard can be modeled as switching to a finer level of probability values. We will look at three examples: 1) the sand case, 2) the lottery case, and 3) the bank case, well-known in the contextualism literature and to be discussed in Section 7.4.

With the sand example and the possibility of counting the individual grains, we saw that whereas some properties may be vague on one level, the region of vagueness may be shifted or removed by going to a finer level. We also indicated that such a change in context is typically triggered by a higher reward for the agent, since the refinement of the level usually requires more effort of the agent (observing for a longer time, using larger equipment, ...).

Let us find an application of this in our lottery example: suppose a person is quick to say “I do not believe that my ticket will win.” Of course, the probability of winning for any ticket is non-zero, so he can never be absolutely certain that he will not win, no matter how large the lottery is and how small the probability. Yet, there is a level at which the winning odds of a single ticket are indistinguishable from zero. In other words, the agent only makes a very small rounding error if he takes his probability of winning to be zero. Doing so may not be justified on the highest standards of rationality, but for practical purposes it seems to be good enough and certainly not completely irrational.

Of course, the agent is by no means forced to ignore the small difference of his winning odds compared to zero. Suppose the same agent notices that he has just been robbed and really needs the money. He may now be more receptive of the very small chance that he might win the lottery with his single ticket, which would solve his current financial problems. This change of context may make him reconsider his previous statement, a process that can be modeled as a transition to a finer level.

Relative analysis may also be considered in relation to contextual identity. As we have seen, Stratified Belief formalizes a probability being ‘sufficiently close’ to unity in LT as that probability being ‘ultraclose’ to 1 on a certain level ($\simeq_V 1$). In other words, it requires the probability to be indistinguishable from 1. Whether or not this is the case, depends on the context level for all values of the probability except for those *exactly* equal to 1. In the philosophical literature, there are some well-known paradoxes related to identity, such as the Ship of Theseus. It has been suggested that there is a contextual element to the identity predicate involved in these paradoxes. Some examples of the context-dependence of words as ‘same’ (identical) and ‘different’ (not identical) are discussed in [7, p. 22–24]. Also Douven and Decock comment on the vagueness and relativity of the identity predicate [11]. We conclude here that the relation \simeq_V can be used to formalize these ideas of contextual identity.

7.3. Relation to the epistemicist account of vagueness

7.3.1. Threshold-free scales of belief

Since largeness is a vague concept, beliefs about topics in which largeness is a crucial element, such as a large lottery, ‘inherit’ some of this vagueness. Therefore, a model for belief about ultralarge lotteries has to be construed within a framework capable of handling vagueness, and, as we have seen, relative analysis provides such a framework. We have used it as the basis for a threshold-free model of rational belief: the model of Stratified Belief.

It may be argued that our model for Stratified Belief—and associated scales of belief—can be interpreted as having ‘hidden’ thresholds too. In that case, however, it should be noted that it is of crucial importance that the thresholds are implicit. Any attempt to make their value explicit makes the stratified analysis core of the model collapse back into standard analysis, which means we get back all the threshold-related problems (such as the total failure of CP).

²⁵ Whereas conditionalizing on zero-probability events is problematic when using Kolmogorov’s ratio formula [20], it is unproblematic to conditionalize on events with a relatively infinitesimal probability, since these are just ordinary, non-zero real numbers. This is similar to the solution to the countably infinite lottery problem given in [46], where the use of NSA allows one to assign (actual) infinitesimal probabilities to single tickets.

So far, we have only discussed yes–no beliefs. Now we return to the topic of qualified belief and subjective probability, quickly passed over in the Introduction. In our language, there are various words and word combinations to express subtle differences in the strength of our beliefs. We may say: ‘I suspect that’, ‘I believe that’, or ‘I am convinced that’ to give only three examples in the order of firmness of belief.²⁶

These various expressions could be thought of as expressing different degrees of belief. In the threshold-based model, a higher degree of belief can be made to correspond to the probability being at least equal to a higher threshold value. In our model of Stratified Belief, we replace the degrees with the threshold-free notion of scales of belief: a higher scale of belief corresponds to the probability being indistinguishable from unity still at a finer level.²⁷

7.3.2. Vague thresholds

We will now show how we can adapt Stratified Belief to encompass weaker forms of belief. This adapted form of SB, SB_θ , can also be used to model the epistemicist account of vagueness [48,49], which assumes that the threshold does have a sharp value, but that this value is essentially inaccessible to us.

Let us fix an approximate threshold θ (in the interval $[\frac{1}{2}, 1]$) and define SB_θ by:²⁸

Definition 3 (Stratified Threshold Belief, SB_θ).

$$B(x) \in R_{\alpha, V} \Leftrightarrow P(x) \succsim_V \theta$$

where $P(x) \succsim_V \theta$ means that either $P(x) > \theta$ or $P(x) \simeq_V \theta$.

Applied to an ultralarge lottery, the Stratified Conjunction Principle would remain valid for two singletons, or for an arbitrary event E_1 and a singleton (but no longer for arbitrary E_2). Observe that the initial formulation of Stratified Belief is a special case of SB_θ with $\theta = 1$.

If we compare SB_θ to the usual threshold-model, we see that although the threshold θ is a specific real number, it functions merely as an arbitrarily chosen representative of all the numbers $r \in \mathbb{R}$ such that $r \simeq_V \theta$.²⁹ The vagueness works in two ways: in determining what the threshold is or ought to be, and in the formation of beliefs. Therefore, SB_θ can be used to model thresholds that have no precise value, as well as thresholds that do have a precise value but which is inaccessible to us. Using the latter interpretation, SB_θ can be thought of as a formal representation of the epistemicist account of vagueness.

7.4. Application of Stratified Belief to similar problems

The model for Stratified Belief was motivated by cases in which information in terms of objective probabilities is explicitly available. It replaces the idea of degrees of belief with scales of belief. However, the former notion is not restricted to cases in which the objective probabilities are known, and actually fits better with cases in which they are not. In this subsection, we apply the analysis in terms of Stratified Belief to two examples in which the exact values of the objective probabilities are not known. Moreover, the first example is usually presented in terms of knowledge rather than (rational) belief. Nevertheless, Stratified Belief does seem to provide a sensible account of this case, too.

In the contextualism debate, there are many examples without explicit probabilities. A popular example is Keith DeRose’s ‘bank case’: given circumstantial evidence, we may claim to know that the bank is open on Saturday if little is at stake, but when the stakes are higher, we may deny doing so. When the stakes are higher, the epistemic standard rises, and this may be modeled by using a finer level. The (implicit) probability value for the bank being open on Saturday is so close to unity that on the coarse level used in a case with low stakes, it is indistinguishable from unity. One may be aware that there is no absolute certainty, but the difference is a relative infinitesimal and therefore inappreciable: it is rational to believe or say one knows the bank to be open. When the stakes are higher, there is an incentive to reconsider the importance of the difference between the relevant probability and unity. Using a finer level, a previously infinitesimal difference becomes appreciable and it becomes rational to say that one does not know that the bank will be open. In the contextualism literature, there are more complicated examples (with different speakers as well as different contexts), but because the second axiom of relative analysis ensures that it can always be established which of two levels is at least as fine as the other, these cases could be analyzed in terms of levels, too.

²⁶ As indicated before, we do not consider the knowledge-version of the Lottery Paradox. Although ‘I know that’ would certainly rank as a very strong expression of firmness of belief, knowledge requires something more than firm belief.

²⁷ It may seem appealing to let knowledge correspond to the finest level. Axiom 3, however, states that for any level there are numbers that are ultrasmall compared to it. Therefore, for any level there is a finer level, and ‘the finest level’ does not exist (an illustration of incompleteness). However, in the case of a specific N -ticket lottery there are levels at which all the relevant knowledge is available: those which contain N and (thus) $\frac{1}{N}$. The point is that there is no such level with this property for general N .

²⁸ Thanks to Karel Hrbacek for this suggestion.

²⁹ Note that these numbers—like a level—do not form a set. In the context of extension-style NSA, this would be called the ‘halo’ of the number θ . Here it is again a predicate, not a set, and level-dependent.

A problem similar to the Lottery Paradox arises in thinking about elections: it is rational for anyone to believe that the impact of his or her single vote on the result of the elections is negligible, because it is only one out of an (ultra)large number of votes. However, the combined impact of all individually negligible votes is not negligible at all. This is called the paradox of (non-)voting [37]. Clearly, the same analysis given here for the ultralarge lottery can be applied to the paradox of voting: the paradox can be blocked by noting that the unrestricted application of mathematical induction would be unreasonable.

8. Conclusions

Our search for a model of rational beliefs based on probabilistic information was motivated by the Lottery Paradox. We observed that vagueness is an essential ingredient to that paradox, that there is a contextual aspect to this vagueness, and that a similar problem exists for infinite lotteries. We also observed that we expect a weakened form of the Conjunction Principle to hold for rational beliefs. The combination of these observations led us to the application of relative analysis. Based on the relative analysis framework, we formulated a soritic, contextual version of the Lockean Thesis. This led us to a new definition of rational belief as ‘almost certainty’ (including absolute certainty), which we called Stratified Belief. The question of whether the Lockean Thesis and our proposed Stratified Belief model for it are psychologically plausible is taken up in a separate paper [45].

We also investigated the aggregation of this type of beliefs. Kyburg’s own response to the Lottery Paradox was to abandon CP [27], whereas many later authors have tried to rescue (part of) it [47]. We found that a weakened version of CP indeed holds for stratified beliefs. Because the aggregation is restricted to ‘a few’ (a standard number of) beliefs, the Lottery Paradox does no longer occur. Based on the lottery example, we may compare aggregating beliefs to doing a calculation based on rounded values: it is better to avoid this, but if the rounded values are all you have, some calculations still give reasonably good outcomes. All of this can be stated more rigorously using the language of relative analysis.

One of the observations that led to our solution, was the analogy with an infinite lottery. In an infinite lottery, it is possible that a specific ticket will win, but the real-valued probability assigned to this possibility is zero, exactly the same as for the impossible event. This leads to some counter-intuitive results (such as the failure of Countable Additivity), which can be blocked using NSA, by allowing infinitesimals as the value of the probability function. In a finite lottery, no matter how large, the probability of any single ticket is strictly larger than zero. However, in our mental representation of it and our resulting rational beliefs, we may not always be able to distinguish between an event with a very small probability and the impossible event. This leads to the counter-intuitive result that even the conjunction of two such rational beliefs is not guaranteed to be rational. Again, NSA can be applied to solve the puzzle: using relative analysis, probabilities that are indistinguishable from zero can be modeled as ultrasmall numbers or relative infinitesimals. As was already mentioned, the concept of beliefs based on this framework does allow the conjunction of, for example, two beliefs.

In short, by looking at the Lockean Thesis from the viewpoint of relative analysis, we have found a new way to incorporate its inherent soritic and contextual nature in the formalization. This leads to a definition for and aggregation rule of rational beliefs, whose combination does not lead to a Lottery Paradox. However, we do not regard dealing with the Lottery Paradox as the end goal of this paper. It serves merely as a case study, prompting reflection about the presence of vagueness in our statements and judgments concerning topics about which we have precise probabilistic information. It is our conviction that the proposed model of Stratified Belief can be a useful tool for epistemology in general. In particular, the model matches well with various contextualist approaches—in epistemology as well as metaphysics—that have appeared in the last few decades.

Apart from this application to rational beliefs, relative analysis may have a further role to play in formal philosophy: it is a powerful mathematical system capable of dealing with vagueness and its contextual aspect. To what extent this is a contribution to the philosophy of vagueness in general may be worth further reflection.

Acknowledgements

I am grateful to Sam Sanders for introducing me to relative analysis and to Karel Hrbacek for the e-mail correspondence on relative analysis and vagueness. For detailed feedback on a previous version of this text, I thank Igor Douven and the members of the Promotion Club Cognitive Patterns in Groningen, in particular David Atkinson, Barteld Kooi, Erik Krabbe, Jeanne Peijnenburg, Bryan Renne, Jan-Willem Romeijn, and Sonja Smets.

References

- [1] P. Achinstein, *The Book of Evidence*, Oxford University Press, 2003.
- [2] E.W. Adams, Aspects of inductive logic, in: J. Hintikka, P. Suppes (Eds.), *Probability and the Logic of Conditionals*, North-Holland, Amsterdam, 1966, pp. 253–316.
- [3] V. Benci, M. Di Nasso, Alpha-theory: an elementary axiomatic for nonstandard analysis, *Expositiones Mathematicae* 21 (2003) 355–386.
- [4] V. Benci, M. Di Nasso, M. Forti, Nonstandard methods and applications in mathematics, in: N.J. Cutland, M. Di Nasso, D.A. Ross (Eds.), *The Eightfold Path to Nonstandard Analysis*, in: *Lecture Notes in Logic*, vol. 25, Association for Symbolic Logic, AK Peters, 2006, pp. 3–44.
- [5] V. Benci, L. Horsten, S. Wenmackers, Non-Archimedean probability, 2012, <http://dx.doi.org/10.1007/s00032-012-0191-x>, *Milan Journal of Mathematics*, in press.

- [6] R. Booth, E. Richter, On revising fuzzy belief bases, *Studia Logica* 80 (2005) 29–61.
- [7] R. Crawshay-Williams, *Methods and Criteria of Reasoning*, 1957, reprinted by: The International Library of Philosophy (2001).
- [8] B. de Finetti, *Theory of Probability*, Wiley, 1974, translated by: A. Machi and A. Smith.
- [9] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, *The Annals of Mathematical Statistics* 38 (1967) 325–339.
- [10] K. DeRose, *The Case for Contextualism: Knowledge, Skepticism, and Context*, Oxford University Press, 2009.
- [11] I. Douven, L. Decock, Identity and similarity, *Philosophical Studies* 151 (2009) 59–78.
- [12] I. Douven, J. Uffink, The preface paradox revisited, *Erkenntnis* 59 (2003) 389–420.
- [13] I. Douven, T. Williamson, Generalizing the lottery paradox, *British Journal for the Philosophy of Science* 57 (2006) 755–779.
- [14] P. Égré, Soritical series and Fisher series, in: A. Hieke, H. Leitgeb (Eds.), *Reduction: Between the Mind and the Brain*, in: Publications of the Austrian Ludwig Wittgenstein Society, vol. 12, Ontos Verlag, 2009, pp. 91–115.
- [15] R. Foley, Justified inconsistent beliefs, *American Philosophical Quarterly* 16 (1979) 247–257.
- [16] R. Foley, Beliefs, degrees of belief, and the Lockean thesis, in: F. Huber, C. Schmidt-Petri (Eds.), *Degrees of Belief*, Springer Science & Business Media B.V., 2009.
- [17] H. Gaifman, Towards a unified concept of probability, in: R.B. Marcus, G.J.W. Dorn, P. Weingartner (Eds.), *Logic, Methodology and Philosophy of Science VII*, in: *Studies in Logic and the Foundations of Mathematics*, vol. 114, Elsevier, Amsterdam, The Netherlands, 1986, pp. 319–350.
- [18] H. Gaifman, Non-standard models in a broader perspective, in: A. Enayat, R. Kossak (Eds.), *Non-standard Models of Arithmetic and Set Theory*, in: *Contemporary Mathematics*, vol. 361, American Mathematical Society, 2004, pp. 1–22.
- [19] H. Gaifman, Vagueness, tolerance and contextual logic, *Synthese* 174 (2010) 5–46.
- [20] A. Hájek, What conditional probability could not be, *Synthese* 137 (2003) 273–323.
- [21] G. Harman, *Change in View: Principles of Reasoning*, MIT Press, Cambridge, MA, 1986.
- [22] J. Hawthorne, L. Bovens, The preface, the lottery, and the logic of belief, *Mind* 108 (1999) 241–264.
- [23] K. Hrbacek, Stratified analysis?, in: I. van den Berg, N. Neves (Eds.), *The Strength of Nonstandard Analysis*, Springer, 2007, pp. 47–63.
- [24] K. Hrbacek, O. Lessmann, R. O'Donovan, Analysis with ultrasmall numbers, *The American Mathematical Monthly* 117 (2010) 801–816.
- [25] P. Klein, The virtues of inconsistency, *Monist* 68 (1985) 105–135.
- [26] A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitrechnung*, *Ergebnisse Der Mathematik*, 1933, translated by N. Morrison, *Foundations of Probability*, 2nd ed., Chelsea Publishing Company, 1956.
- [27] H.E.J. Kyburg, *Probability and the Logic of Rational Belief*, Middletown, 1961.
- [28] A. Lang, The limited capacity model of mediated message processing, *Journal of Communication* 50 (2000) 46–70.
- [29] S. Lavine, Finite mathematics, *Synthese* 103 (1995) 389–420.
- [30] H. Leitgeb, *Reducing belief simpliciter to degrees of belief*, 2010, unpublished manuscript.
- [31] D. Lewis, A subjectivist's guide to objective chance, in: R.C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability*, University of California Press, 1980, pp. 263–293.
- [32] O. Magidor, Strict finitism and the happy sorites, *Journal of Philosophical Logic* 41 (2012) 471–491.
- [33] D. Makinson, The paradox of the preface, *Analysis* 25 (1965) 205–207.
- [34] H.L. Mencken, *Prejudices: First Series*, 1919, reprinted by: Wildside Press LLC, 2005.
- [35] D.K. Nelkin, The lottery paradox, knowledge, and rationality, *The Philosophical Review* 109 (2000) 373–409.
- [36] E. Nelson, Internal set theory: a new approach to nonstandard analysis, *Bulletin of the American Mathematical Society* 83 (1977) 1165–1198.
- [37] G. Owen, B. Grofman, To vote or not to vote: the paradox of nonvoting, *Public Choice* 42 (1984) 311–325.
- [38] P. Péraire, G. Wallet, Une théorie relative des ensembles internes, *Comptes rendus de l'Académie des sciences. Série, Mathématique* 1 (1989) 301–304, 308.
- [39] L. Price, *Dialogues of Alfred North Whitehead*, Little Brown, Boston, MA, 1954, reprinted by David R. Godine Publisher, Boston, MA, 2001.
- [40] G. Priest, What is so bad about contradictions?, *The Journal of Philosophy* 95 (1998) 410–426.
- [41] H. Putnam, *Reason, Truth, and History*, Cambridge University Press, 1981.
- [42] F.P. Ramsey, Truth and probability, in: F.P. Ramsey (Ed.), *The Foundations of Mathematics and Other Logical Essays*, Harcourt, Brace, 1931, original paper from 1926.
- [43] A. Robinson, *Non-Standard Analysis*, North-Holland Publishing Company, 1966.
- [44] L.G. Shapiro, G.C. Stockman, *Computer Vision*, Prentice Hall, 2002.
- [45] S. Wenmackers, The Lockean Thesis and Stratified Belief: Rational beliefs for real people?, 2013, in preparation.
- [46] S. Wenmackers, L. Horsten, Fair infinite lotteries, *Synthese* 190 (2013) 37–61.
- [47] G. Wheeler, A review of the lottery paradox, in: W.L. Harper, G. Wheeler (Eds.), *Probability and Inference: Essays in Honour of Henry E. Kyburg, Jr.*, King's College Publications, 2007, pp. 1–31.
- [48] T. Williamson, *Vagueness*, Routledge, 1994.
- [49] T. Williamson, Imagination, stipulation and vagueness, *Philosophical Issues* 8 (1997) 215–228.